G. Leffers

ENTERED

#9
S.Q.
8/24/00

```
        RAW SEQUENCE LISTING                 DATE: 08/15/2000
        PATENT APPLICATION:  US/09/430,590A   TIME: 11:58:10

        Input Set : A:\ES.txt
        Output Set: N:\CRF3\08152000\I430590A.raw
```

RECEIVED

AUG 22 2000

TECH CENTER 1600/2900

```
   3 <110> APPLICANT: POULTER et al.
   4        POULTER et al.
   6 <120> TITLE OF INVENTION: UNUSUAL RETROTRANSPOSON FROM THE YEAST CANDIDA ALBICANS
   8 <130> FILE REFERENCE: 674521-2001.1
  10 <140> CURRENT APPLICATION NUMBER: 09/430,590A
  11 <141> CURRENT FILING DATE: 1999-10-29
  13 <160> NUMBER OF SEQ ID NOS: 3
  15 <170> SOFTWARE: PatentIn Ver. 2.1
  17 <210> SEQ ID NO: 1
  18 <211> LENGTH: 6426
  19 <212> TYPE: DNA
  20 <213> ORGANISM: Candida albicans
  22 <400> SEQUENCE: 1
  23 tgttggtttg tgcactattt tgtgtcagaa actgatcaat gaaaatgatg gttattatga 60
  24 gaatggaaaa tttttccatc acacatcagg tgatgacaga actaaactat attgtgtagt 120
  25 ataaataagg gtatgaaata ccaacatccc agaatatcaa cgagatagaa gggaggagtt 180
  26 tcaatatata tcttgtgaat aataacttcg ttctaattca ctatacacaa ctagacgtgt 240
  27 acacgctcaa tctcaggtaa agaaagttta tattccatca gattagaagt cgatagtgat 300
  28 aatcatttcg tcccaaatta gcgttgtata aattcagtcc tcagatttgt attattgatt 360
  29 gatagtttcg aagtttgaag gtacagaatt tcacaagatg agttccgcaa agaatgatga 420
  30 taacgaaggg aaggtcatgg aaagtgttga tcaagctaat gctattagta aggtggatga 480
  31 acatatcaag gctagattca atatgctttt cataaaattt aatgacttac ctaagttggc 540
  32 cgtcggtaat cagaaaagcg tggataaatg gaatgaagaa tttaaatatt tccacgttgc 600
  33 tttatatgtt gatgatattc ttatggttgg aagttcacaa aaagttattg ataatttgt 660
  34 ttaccccgat gtttttggaat ttttgcttga ctataatcct aaagataaat tcaaggttaa 720
  35 aaaggtagaa ggtatttatt ttactggttg gtgtttacaa atgtgtttac agtccatttt 780
  36 tgataggttc agattgatca tgattctaa gctaccaaag cacttgcaaa aggaagcaaa 840
  37 cttaatcaaa gctgcttatg atgctgttac taaatctaaa gattatacca ttactagtaa 900
  38 gatcttgctg aagtttgtaa acgttgaaca tgagttagtg gtttgctata accttccata 960
  39 tttgctgcag gtggaagaga aacttgagga aatactctac aacacttcaa acgttgtcga 1020
  40 tgagtatgtc cgtagtcttc caaatctcat aggtcaagtc ttgtacttca atcatgtgaa 1080
  41 gaaatcagag gctttaagtt tgttttttgaa tattcatgcc tcatactact caaagtggat 1140
  42 tcaagctgac aatgatacat cagtactccc aagttgctct accatagctg aagaaatgtg 1200
  43 tgatcatcct gattatgcta gattggttga cattccaagc aacaaatatg aacttaatct 1260
  44 tattgttagt ttaccagcac cagagaaacc aaaaggaaaa ccagaggaga actcactgga 1320
  45 acaatctcaa aagaagaacc tgaaatcaag aaagagaaat aagaaacatc caaaatcaga 1380
  46 taacgataaa ggtgaaaaag aaaaagaaaa agaaaaaact tcactggaat gaaaaacagg 1440
  47 tgctgcttct attaattgtg taatgaatat acataattgc agcaaaacca cgtttccagt 1500
  48 agaaaattct cattctctta atgcttcttt gaacgtaatg aattttaaag gttaaggtt 1560
  49 taacaagtat ctagtgtatg atactggtgc cacaatatct gttgtgaaca ataaagatat 1620
  50 attgctgaat gttaaggacg caacaattga agtttctgtt gctgatggtg ctacattaga 1680
  51 agcagattgt attggtgatc taattatcag agtcggtatt gtctcgatta cgttagaaga 1740
  52 tacattgtat ttaccagaaa gttcctttaa tcttgtgagt ttgaaacaaa ttgaagaacg 1800
  53 aggatttaat gttcttatta ctaaagaatc agtgattgta tttaaccaaa atgtggctcc 1860
  54 tactattatt gcttcaagga agaatgctgc tgatctttat atgggtcctc aattcagtga 1920
  55 agaatcttta gaatgtgatt ttgattatga tggtttggca gatatgttgt ccaatgctaa 1980
  56 ccaagatgac aaagataaat caagtatgaa tgaaatgtca gaatatcaag aacatgatta 2040
```

```
 57 tagttctcga gcattaataa attctttgac ggaggttgat gttttagatg ttgaaatttc 2100
 58 cccatatgga gttgaacaat tgctaccaac tggagataag aacgatattt ataatttcca 2160
 59 tttgatgtca aatcatatgt ccattgagaa aatcttgttg ttacaaaaat accagggtct 2220
 60 cgtacttcac acttcaaaag agagtcttca aaagattgct gattgtaagg tatgtctatt 2280
 61 atcgaatgcc aaacagagaa gtcacaatca tcattcagaa agaaaagcct cgagaagaca 2340
 62 tgagagactt cattgtgata ctctcggtcc atttaggtcc gaaaataaca agtggtattt 2400
 63 aacgtctgtt atagatgaac atacgggtta cattgaagga attattacta aagacagaaa 2460
 64 ggtaaaggat ctcttaattc aacgattaaa gatctggaat aatcggttta acgataaggt 2520
 65 ggcatacttc agaagtgata atgctcctga gttcccacaa ccttctgatt tagctgagtt 2580
 66 cggtatttgg agggagacta tagcggcata tctgcctgag cttaatggtc tcgccgaggt 2640
 67 tgttaataaa ttgattttac aacagattta caggatcgtt gtgacacttg gtccacaaat 2700
 68 actcaagttg atttattatg tgattcaata ttctattaca atgatcaacc acactccacg 2760
 69 tcgttcactc aagggacaaa cccccttatg ttgctattat caattaagtg agggaaattt 2820
 70 ctaccggttt ccttttgcca tcgattgtgt cgttacattt agtaatgcca tcgaaaagaa 2880
 71 ccgttacgga gttacatcaa ctaaaggagc tccttcatcg atcatggggtg ctgtgattgg 2940
 72 ctacgctagc gattgtttta gttattacgt gttgctaaaa aatatgcggt gtgatattat 3000
 73 ccttagccct aatgtccgta tattgcgaag ctatgaggtt attaactcct atctcaaaaa 3060
 74 cttatccact acacctatgt cacacattgt tcctatggct gaaggtatcc agggaaggca 3120
 75 actgggcgct cagtacgagg tacgcggaac atatgtggaa agtgaatatg acaatacaaa 3180
 76 tgacgtgatg cacatgccca aagagtcata ttcagttcag ccagcatcgt ttactttaac 3240
 77 tacgggtaac agttctaacg aatatgttat aaatgatgat ccagtacaga ttaccattga 3300
 78 gaatcccgat gattttcta accctcttca actaactgaa gaatcacacg atatggtatc 3360
 79 cgaagtaaaa tcggatgaga atcctaaacc cagtctccac gagctaacac ctggggataa 3420
 80 tccggtgtct aaacctcctc aacttggtac cgagacttca gtaataggga agtctaaaga 3480
 81 gcctattaca aaccacacaa aggacgcccc ttccatccag gggagggacc ataaacgcct 3540
 82 ggaatctact gctcaggttg gactatcaca ccaacccag actggtactc ccgcttcgga 3600
 83 ggagtcaaaa ttgtcaggaa cagatcattt cggtgtcgac gttgttaaag aaacagtctc 3660
 84 agaagattgg catacttctg actacccaga aactagtgct gaagatgaac agcaaaatcc 3720
 85 ctcgttactg gctaataaga atcgggtaac tgaaaaaata gatgagggag aaaatatttc 3780
 86 atttccggg ggtgatgatg attctgtcgt gatcaactca aatgttgagc aatctaatgt 3840
 87 tgaaacagag gatgctggta acagtccaat tcaagacgaa gtttctcaag agggaagaat 3900
 88 acttaatgaa caaactgata tagttgatac tgttgctaaa gttattgaga atgaaaaaat 3960
 89 ctctcctatt aattcattag atgatcatac tgaacttgct acagactcgg gaaatgatag 4020
 90 caattcaaca gaatccgaca ttcaatcgaa aaatgaaata tcaccagtga ttaatgagaa 4080
 91 aaatactgaa ataatccaaa aacacattga aagtatcctt gctgataaga gattggatga 4140
 92 atttgaaacg tataatgttg atgaaattga gaatgtgatt aatgacgatg acattgctga 4200
 93 agctaatcca ctaccagatg aaaataatga tgttcagatg aatgagagtt ttgataataa 4260
 94 tcatagcatg tcacgagcaa agaagaaata cacatttgag aaagaagtta acgaaaaaat 4320
 95 tgctggtact aaacattcac ttgatacaac tgatccaaga gaagcaatca gagtgttaaa 4380
 96 tactggtgaa accaagagaa tcgaacccaa gaaaagaagg gtgcctatca ctgtgaaatt 4440
 97 aaacaaaaga tcgcaataca agtcaccata tgttacaaga agtggtagaa cggttataaa 4500
 98 ccccaagagg tatttacatg cggtcgtcaa caaaatcgac tataatgatc cgggatggat 4560
 99 aaagtcaatg aatgctgaac tagagaaatt tagatcaaaa gatgtttacg aagaagttcc 4620
100 aattcccacc ggtgtgaagc ctatatctat gggttgggta catactgaga aaattgattc 4680
101 tctcaaaggt gttgttcgga aatcacgttg tgttgtccat ggcaacgac aaaaggaaaa 4740
102 attggattat gacccttta gtgttagttc acctgttata gatcttgtga ctataagatt 4800
103 attgacaata ataggttgtg aattaggaat gacaattcaa catttagacg tcgagtcggc 4860
104 gtatctaaat gcctctatta ctcattcaaa tccaatttat gtctttcctc ctaaatcagt 4920
105 acctttgaag aaaaaccatt gttggttatt gaaacgttct gtctatgggt taaaacagtc 4980
```

RAW SEQUENCE LISTING   DATE: 08/15/2000
PATENT APPLICATION: US/09/430,590A   TIME: 11:58:10

Input Set : A:\ES.txt
Output Set: N:\CRF3\08152000\I430590A.raw

```
106 gggtttggaa tggtatcaca ctatcaaaag agtattggaa gacattggtt ttactcaagt 5040
107 tttatatgtt gatgatattc ttatggttgg aagttcacaa aaagtattg ataattttgt 5100
108 ggatcaattg agagatcatt ttgaagttaa agtgtttggt gaaatatcaa attatcttgg 5160
109 tattgaattt cgtaaaaccg aatctggtta tattttatct caagaaaaat ttctcaagaa 5220
110 attacttaag gatttcaaac tagatgactc atatgggaaa aacataccct ggattccgaa 5280
111 tgacaaatat gaaaaggttg caataattcg tgaaaacgtt aatccagaga atgattttga 5340
112 aaaggttccg aatgagacat tgcttgaccc tgatgctaaa aaactatacc aaagtggtgt 5400
113 tggcctgctt ttatgggctg ccacaaacac acgtccagat atatcggtcg tagtgaattc 5460
114 gttgggttct aaatctgcaa atccaaatgt ccatgattat gagaaattga tttattgtct 5520
115 taggtatatc aaaaatagca tgggatatca cattgagtac aaaagaaaca gattgaatat 5580
116 accaccaaaa tcatttgtta tcgaatgttt cagtgatgcg tcatttgcac caggattgga 5640
117 tagaaaatct attagtggaa ctttgattta tgtgaatgga aatttggtgc aatgggcgac 5700
118 caaaaaacaa acggtcatag cacaaagctc agcagcttgt gaaatgttgg ctctaaatta 5760
119 tacaatgttg aaagctatcg aaataaaaaa ccatttaatg gatttgggtt ttgaagtagg 5820
120 taagatacat tgtcatcaag acaaccaagc tgtgattaaa gttttgagaa ataactattg 5880
121 tcacccacat cgaccaatag atatctgcta taagtttcta cgccaattga tcaatgataa 5940
122 agtattttca atatcctatg tgaagacaaa tgataattac gccgattgta tgactaagtg 6000
123 tctaagtcgt gctaaattca aagcattcgt tgagggtatg ataaaacggt tagacctaga 6060
124 agataatcaa acactgatac aaaatgcaat aacggcagaa taagtggatt tatcattact 6120
125 attatcgtaa tgctcaatca ggggagtgtt ggtttgtgca ctattttgtg tcagaaactg 6180
126 atcaatgaaa atgatggtta ttatgagaat ggaaaatttt tccatcacac atcaggtgat 6240
127 gacagaacta aactatattg tgtagtataa ataagggtat gaaataccaa catcccagaa 6300
128 tatcaacgag atagaaggga ggagtttcaa tatatatctt gtgaataata acttcgttct 6360
129 aattcactat acacaactag acgtgtacac gctcaatctc aggtaaagaa agtttatatt 6420
130 ccatca                                                        6426
```

133 <210> SEQ ID NO: 2
134 <211> LENGTH: 324
135 <212> TYPE: PRT
136 <213> ORGANISM: Candida albicans
138 <400> SEQUENCE: 2

```
139 Met Ser Ser Ala Lys Asn Asp Asp Asn Glu Gly Lys Val Met Glu Ser
140   1               5                  10                  15
142 Val Asp Gln Ala Asn Ala Ile Ser Lys Val Asp Glu His Ile Lys Ala
143              20                  25                  30
145 Arg Phe Asn Met Leu Phe Ile Lys Phe Asn Asp Leu Pro Lys Leu Ala
146          35                  40                  45
148 Val Gly Asn Gln Lys Ser Val Asp Lys Trp Asn Glu Glu Phe Lys Tyr
149      50                  55                  60
151 Phe His Val Ala Tyr Pro Asp Val Leu Glu Phe Leu Leu Asp Tyr Asn
152  65                  70                  75                  80
154 Pro Lys Asp Lys Phe Lys Val Lys Lys Val Glu Gly Ile Tyr Phe Thr
155              85                  90                  95
157 Gly Trp Cys Leu Gln Met Cys Leu Gln Ser Ile Phe Asp Arg Phe Arg
158          100                 105                 110
160 Leu Ile Met Ile Ser Lys Leu Pro Lys His Leu Gln Lys Glu Ala Asn
161          115                 120                 125
163 Leu Ile Lys Ala Ala Tyr Asp Ala Val Thr Lys Ser Lys Asp Tyr Thr
164      130                 135                 140
166 Ile Thr Ser Lys Ile Leu Ser Lys Phe Val Asn Val Glu His Glu Leu
```

```
167 145                  150                 155                 160
169 Val Val Cys Tyr Asn Leu Pro Tyr Leu Ser Gln Val Glu Glu Lys Leu
170 170                 165                 170                 175
172 Glu Glu Ile Leu Tyr Asn Thr Ser Asn Val Val Asp Glu Tyr Val Arg
173             180                 185                 190
175 Ser Leu Pro Asn Leu Ile Gly Gln Val Leu Tyr Phe Asn His Val Lys
176         195                 200                 205
178 Lys Ser Glu Ala Leu Ser Leu Phe Leu Asn Ile His Ala Ser Tyr Tyr
179     210                 215                 220
181 Ser Lys Trp Ile Gln Ala Asp Asn Asp Thr Ser Val Leu Pro Ser Cys
182 225                 230                 235                 240
184 Ser Thr Ile Ala Glu Glu Met Cys Asp His Pro Asp Tyr Ala Arg Leu
185             245                 250                 255
187 Val Asp Ile Pro Ser Asn Lys Tyr Glu Leu Asn Leu Ile Val Ser Leu
188             260                 265                 270
190 Pro Ala Pro Glu Lys Pro Lys Gly Lys Pro Glu Glu Asn Ser Ser Glu
191         275                 280                 285
193 Gln Ser Gln Lys Lys Asn Ser Lys Ser Arg Lys Arg Asn Lys Lys His
194     290                 295                 300
196 Pro Lys Ser Asp Asn Asp Lys Gly Glu Lys Glu Lys Glu Lys Glu Lys
197 305                 310                 315                 320
199 Thr Ser Ser Glu
203 <210> SEQ ID NO: 3
204 <211> LENGTH: 1576
205 <212> TYPE: PRT
206 <213> ORGANISM: Candida albicans
208 <400> SEQUENCE: 3
209 Lys Thr Gly Ala Ala Ser Ile Asn Cys Val Met Asn Ile His Asn Cys
210     1               5                 10                 15
212 Ser Lys Thr Thr Phe Pro Val Glu Asn Ser His Ser Leu Asn Ala Ser
213             20                 25                 30
215 Leu Asn Val Met Asn Phe Lys Gly Leu Arg Phe Asn Lys Tyr Leu Val
216         35                 40                 45
218 Tyr Asp Thr Gly Ala Thr Ile Ser Val Val Asn Asn Lys Asp Ile Leu
219     50                 55                 60
221 Ser Asn Val Lys Asp Ala Thr Ile Glu Val Ser Val Ala Asp Gly Ala
222 65                 70                 75                 80
224 Thr Leu Glu Ala Asp Cys Ile Gly Asp Leu Ile Ile Arg Val Gly Ile
225             85                 90                 95
227 Val Ser Ile Thr Leu Glu Asn Thr Leu Tyr Leu Pro Glu Ser Ser Phe
228             100                 105                 110
230 Asn Leu Val Ser Leu Lys Gln Ile Glu Glu Arg Gly Phe Asn Val Leu
231         115                 120                 125
233 Ile Thr Lys Glu Ser Val Ile Val Phe Asn Gln Asn Val Ala Pro Thr
234     130                 135                 140
236 Ile Ile Ala Ser Arg Lys Asn Ala Ala Asp Leu Tyr Met Gly Pro Gln
237 145                 150                 155                 160
239 Phe Ser Glu Glu Ser Leu Glu Cys Asp Phe Asp Tyr Asp Gly Leu Ala
240             165                 170                 175
```

RAW SEQUENCE LISTING   DATE: 08/15/2000
PATENT APPLICATION: US/09/430,590A   TIME: 11:58:10

Input Set : A:\ES.txt
Output Set: N:\CRF3\08152000\I430590A.raw

```
242 Asp Met Leu Ser Asn Ala Asn Gln Asp Asp Lys Asp Lys Ser Ser Met
243         180                 185                 190
245 Asn Glu Met Ser Glu Tyr Gln Glu His Asp Tyr Ser Ser Arg Ala Leu
246         195                 200                 205
248 Ile Asn Ser Leu Thr Glu Val Asp Val Leu Asp Val Glu Ile Ser Pro
249     210                 215                 220
251 Tyr Gly Val Glu Gln Leu Leu Pro Thr Gly Asp Lys Asn Asp Ile Tyr
252 225                 230                 235                 240
254 Asn Phe His Leu Met Ser Asn His Met Ser Ile Glu Lys Ile Leu Leu
255             245                 250                 255
257 Leu Gln Lys Tyr Gln Gly Leu Val Leu His Thr Ser Lys Glu Ser Leu
258         260                 265                 270
260 Gln Lys Ile Ala Asp Cys Lys Val Cys Leu Leu Ser Asn Ala Lys Gln
261     275                 280                 285
263 Arg Ser His Asn His His Ser Glu Arg Lys Ala Ser Arg Arg His Glu
264     290                 295                 300
266 Arg Leu His Cys Asp Thr Leu Gly Pro Phe Arg Ser Glu Asn Asn Lys
267 305                 310                 315                 320
269 Trp Tyr Leu Thr Ser Val Ile Asp Glu His Thr Gly Tyr Ile Glu Gly
270             325                 330                 335
272 Ile Ile Thr Lys Asp Arg Lys Val Lys Asp Leu Leu Ile Gln Arg Leu
273         340                 345                 350
275 Lys Ile Trp Asn Asn Arg Phe Asn Asp Lys Val Ala Tyr Phe Arg Ser
276     355                 360                 365
278 Asp Asn Ala Pro Glu Phe Pro Gln Pro Ser Asp Leu Ala Glu Phe Gly
279     370                 375                 380
281 Ile Trp Arg Glu Thr Ile Ala Ala Tyr Ser Pro Glu Leu Asn Gly Leu
282 385                 390                 395                 400
284 Ala Glu Val Val Asn Lys Leu Ile Leu Gln Gln Ile Tyr Arg Ile Val
285             405                 410                 415
287 Val Thr Leu Gly Pro Gln Ile Leu Lys Leu Ile Tyr Tyr Val Ile Gln
288         420                 425                 430
290 Tyr Ser Ile Thr Met Ile Asn His Thr Pro Arg Arg Ser Leu Lys Gly
291         435                 440                 445
293 Gln Thr Pro Tyr Gly Cys Tyr Tyr Gln Leu Ser Glu Gly Asn Phe Tyr
294     450                 455                 460
296 Arg Phe Pro Phe Ala Ile Asp Cys Val Val Thr Phe Ser Asn Ala Ile
297 465                 470                 475                 480
299 Glu Lys Asn Arg Tyr Gly Val Thr Ser Thr Lys Gly Ala Pro Ser Ser
300             485                 490                 495
302 Ile Met Gly Ala Val Ile Gly Tyr Ala Ser Asp Cys Phe Ser Tyr Tyr
303         500                 505                 510
305 Val Leu Leu Lys Asn Met Arg Cys Asp Ile Ile Leu Ser Pro Asn Val
306         515                 520                 525
308 Arg Ile Leu Arg Ser Tyr Glu Val Ile Asn Ser Tyr Leu Lys Asn Leu
309     530                 535                 540
311 Ser Thr Thr Pro Met Ser His Ile Val Pro Met Ala Glu Gly Ile Gln
312 545                 550                 555                 560
314 Gly Arg Gln Ser Gly Ala Gln Tyr Glu Val Arg Gly Thr Tyr Val Glu
```

VERIFICATION SUMMARY   DATE: 08/15/2000
PATENT APPLICATION: US/09/430,590A   TIME: 11:58:11

Input Set : A:\ES.txt
Output Set: N:\CRF3\08152000\I430590A.raw